

REMARKS

Claims 1-45 are pending in this application, of which Claims 1, 23 and 45 are independent claims. Claims 1, 2, 4, 7, 15, 19-21 and 23-45 have been amended to correct minor typographical errors, and to define still more clearly what Applicant regards as her invention. A substitute specification is in preparation and will be submitted shortly.

Claims 1-45 were rejected under 35 U.S.C. § 112, second paragraph, as being indefinite, and were also rejected under 35 U.S.C. § 102(a) as being anticipated by "Microsoft Office 97 User Manual (1998)" (Winter et al.).

As is discussed in more detail in the present application, the present invention is intended to address certain inconveniences encountered by users of web pages. As is well known both by those who construct web content and by consumers of web content, the widespread mark-up languages used for web-page design permit the designer tremendous flexibility as to what is included within a single page, as well as its arrangement. The way the designer finds convenient to arrange that content, however, does not necessarily coincide with the consumer's interests, which obviously may be directed to just a portion of the content. The present invention provides means by which it is possible to analyze content of a web page and to reorganize portions of that data. In particular, according to the present independent claims, a plurality of segments which occur in a table, and each of which is pinched between a start tag and a corresponding end tag, are used by dividing the table.

Initially, the claims have been carefully reviewed and amended as deemed necessary to ensure that they comply fully with the requirements of 35 U.S.C. § 112, with

special attention to the points raised in paragraphs 3-6 of the Office Action. Each of the terms mentioned in those paragraphs is believed to be clear, when the claims are read, as they are required to be, after a fully reading of the specification. That is, it is not necessary for the meaning of each claim term to be fully clear from the claims alone, since it may be the case that an inventor has to use language in an unconventional way to describe an invention, which by definition encompasses something new. Accordingly, it has long been established patent law that the clarity and definiteness of the claims is to be determined, not by a review of the claim language alone, but by reading the claims after a full reading, and in light of, the specification, even though the claim language is not limited in scope to the details of the embodiments described in the specification (unless the disclosure itself makes clear to the contrary), and even though limitations are not to be read into the claim language from the details of the specification.

In this instance, Applicant notes that the use of an array of measure to represent a degree of similarity of two or more objects, and the terming of such array "vector", are well known. Examples of such vectors are given in the present application, at page 10, line 24, through page 14, line 4. Similarly, examples of table types relevant to the claims, are given at page 14, line 5, through page 15, line 17. Explanation of possible cut directions, when it is determined that a particular table is to be cut along a predetermined direction (e.g., along a vertical line dividing two columns or along a horizontal one, dividing two rows) is given at page 15, line 21, through page 18, line 14. Comparison of lines of text (relating to "partial line character") is explained at page 36, line 4, through page 37, line 13.

The term "syntax" is discussed briefly at page 40, lines 5-23. That word, as a term to refer to the precise ways in which commands, variable names and the like (to give a few examples) are permitted to be combined with each other according to a programming language, has been well established for a very long time; use of the same term, to describe the manner of fitting together of the various elements used in a given body of data, whether of words in a text or others, is a natural one. The broad term is used in the claims because it is not intended to limit the scope of the claims in respect of a particular type of syntactic analysis to be used, nor the particular kinds of syntactic phenomena to be examined. That a claim is broad, of course, is not the same as the claim being indefinite, and it is not believed that the use of this term renders any of the present claims indefinite.

At page 49, line 22, through page 51, line 1, is a brief discussion of removal of data that appears in conjunction with a table ("supplementary data removal"). Figs. 29A - 29E and the accompanying text explain "analyzing structure of a table" and "regularity of information description".

Withdrawal of the rejection under Section 112 is respectfully requested.

Independent Claim 1 is directed to a document segmentation apparatus that comprises table analyzing means for generating cell position data indicating a positional relationship between cells and cell vectors representing characteristics of the cells, by analyzing a table pinched between a start tag and an end tag in a document to be processed. Table type judging means judge a table type with reference to the cell position data and the cell vectors generated in this fashion, and first segment generating means generate a plurality of segments each of which is pinched between the start tag and the end tag, by

dividing the table with a first method in a case in which the table type is a list type. Also provided are second segment generating means, which generate a plurality of segments each of which is pinched between the start tag and the end tag by dividing the table with a second method in a case in which the table type is a layout type.

Independent Claims 23 and 45 are, respectively, a method and a memory-medium claim corresponding to apparatus Claim 1.

As noted above, in the present invention as set out in the independent claims, a plurality of segments each of which is pinched between a start tag and a corresponding end tag are produced by dividing the table. In contrast, in Excel of Microsoft, a cell can merely be divided by cutting and pasting. Applicant submits that nothing has been found, or pointed out, in the cited art that would teach to suggest a segment division in which a plurality of segments each of which is pinched between a start tag and a corresponding end tag by dividing a table. For at least that reason, Claims 1, 23 and 45 are each deemed to be clearly allowable over the Microsoft Office 97 User Manual.

A review of the other art of record has failed to reveal anything which, in Applicant's opinion, would remedy the deficiencies of the art discussed above, as a reference against the independent claims herein. Those claims are therefore believed patentable over the art of record.

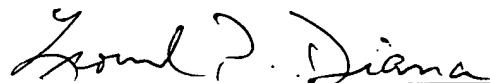
The other claims in this application are each dependent from one or another of the independent claims discussed above and are therefore believed patentable for the same reasons. Since each dependent claim is also deemed to define an additional aspect of

the invention, however, the individual reconsideration of the patentability of each on its own merits is respectfully requested.

In view of the foregoing amendments and remarks, Applicant respectfully requests favorable reconsideration and passage to issue of the present application.

Applicant's undersigned attorney may be reached in our New York office by telephone at (212) 218-2100. All correspondence should continue to be directed to our below listed address.

Respectfully submitted,

  
\_\_\_\_\_  
Attorney for Applicant

Registration No. 29,296

FITZPATRICK, CELLA, HARPER & SCINTO  
30 Rockefeller Plaza  
New York, New York 10112-3801  
Facsimile: (212) 218-2200

NY\_MAIN 393813 v1